The Study on the Identity Verification between the on-Site Face Image and the ID Photo

Zhiguo Yan, Chun Pan and Zheng Xu
The Centre of Internet of Things, The Third Research Institute of Ministry of Public Security
339 Bisheng Road, Shanghai, China

Keywords: Identity Verification, Super-resolution Reconstruction，Deep Learning，Face Verification，RFID.

Abstract: Nowadays, how to pass the security-check passageway with self-service under high throughput is an emerging application challenge. The identity verification technique is the key factor to solve this dilemma, especially in the railway station, bus station and airport etc. The essence of this question is the one vs. one face verification. It involves two crucial application knot, the real-time super-resolution image reconstruction and the real-time face detection and recognition in the video stream under the surveillance scene. To improve the performance of the identity verifications system based on face similarity assessment ， we exploited the deep learning mechanism to train the face detection module and the to realize the super-resolution construction remarkably. The experiment proves its effectiveness.

1 INTRODUCTION

In recent years, with the boosting demand on security check in railway station, bus station and air-port, the self-service passenger pass as an important pre-check mechanism has attract the wide attention. The general application mode lies on the identity verification between the passenger on-site face image and the electronic identification photo. While passenger approached the security-check gate, they will swipe the RFID ID-card and the RFID reader capture the stored electronic photo. At the same time, their face images are captured and stored by the on-the-spot surveillance camera mounted on the security gate. The identity verification system will automatically judge whether the passenger is identical with his carried ID-card based on the face similarity measurement. To those passengers whose score meet the threshold condition, they are allowed to pass the security-check gate. Otherwise they will be blocked.

This procedure involves two key techniques: the identification photo enhancement and the dynamic face verification based on the real-time video stream. As it is well known that the size of the electronic photo stored in the RFID ID-card is 126*102 pixels. This image quality is too weak and unsuitable to execute the face verification. As a premise, the enhancement is a necessary preprocessing. The advisable preprocessing for image enhancement will effectively reduce the false alarm dramaticlly. Considering the low resolution is the key difficulty which impede the practical application, we focus our attention on super-resolution (SR) reconstruction (Dong, 2015) during the image enhancement stage. Different form the work, we adopt the online super-resolution reconstruction for the captured electronic photos. The reason lies in that we can't get the electronic photos prior to the identity verification. Furthermore, with respect to the on-line super-resolution construction for the electronic identification photos, what we can utilized is just themselves. In other words, no redundant information derived from other photos will introduced into the super-construction process.

As the real-time video stream is concerned, the content analysis focus on the pedestrian detection and face region segmentation. The frontal face capture is the crucial step during the dynamic face verification. As to the frontal face image capture, our previous work (Yan, 2013) has addressed a kind of promising methods.

Pedestrian detection in complex scenes is a tough problem. In the general surveillance scene, the unsuitable illumination intensity, the body occlusion and the backlight and shadow exist frequently and give severe adverse impact on face region segmentation. To speed up the pedestrian detection, we execute the down-sampling for those raw frames

Pan C., Xu Z. and Yan Z.
The Study on the Identity Verification between the on-Site Face Image and the ID Photo.
DOI: 10.5220/0006020401300133
In *Proceedings of the Information Science and Management Engineering III (ISME 2015)*, pages 130-133
ISBN: 978-989-758-163-2

The Study on the Identity Verification between the on-Site Face Image and the ID Photo

with 1920*1080 pixels and the DPM algorithm is utilized to find out the pedestrian. In some case, there exists some pedestrians in one frame image which give rise to the multi-face detection. To avoid this phenomena, we tune the focus and the Tele-Wide button to select the advisable view coverage. Moreover, the narrow passageway will limit the occurence of the multi-face in one frame. To the worst case, there still exists multi-face in one frame, we adopt the face with maximum face region as the analyzed target.

The remaining parts of this paper are organized as follows. Section Π gives a simple introduction to the proposed identity verification procedure, and describes super-resolution reconstruction, dynamic face verification. The experimental result is given in section Ⅲ. Conclusions are figured out in the section Ⅳ.

2 THE PROPOSED ARCHITECTURE

As shown in Fig. 1, the identity verification system is composed of four parts: the on-site surveillance camera module, the RFID card reader module, the online face verification module and the automatic security gate control module.

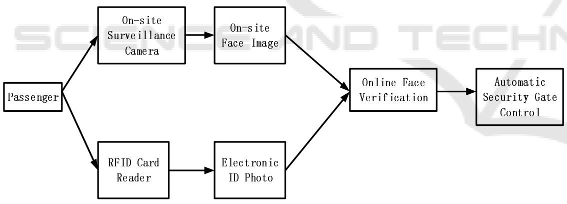

Figure 1: The Schematic Drawing of Identity Verification.

Figure 2: The on-the-spot Deployment of Identity Verification based on Face Verification Mechanism.

In Fig.1, the on-site surveillance camera module is designed to capture the corresponding on-site face image while the passenger is swiping his ID card. The RFID reader module provide the electronic ID photo for online face verification module. From Fig.1 we can figure out that the opening of the security gate is trigged by the positive verification result.

Fig.3 shows the configuration of the identity verification system integrated with luggage X-ray security check device. While passengers approach the security check system, their luggage bags will be transmitted by the transmission belt in the X-ray scanner and they will pass through the security-check gate with self-service. The security system will determine whether allow the passenger to pass the gate according to the joint judgement of luggage security check and the identity verification.

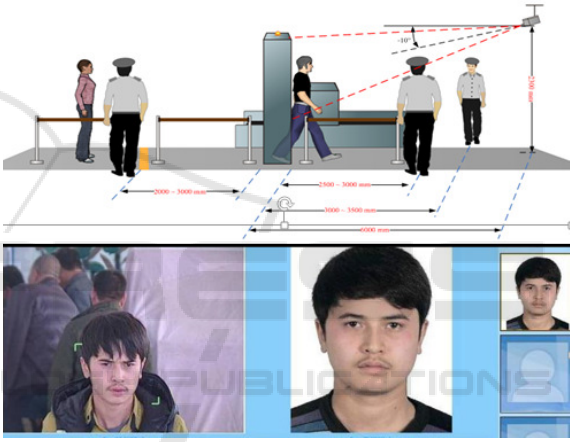

Figure 3: The integration of the Identity Verification Security check gate with the X-ray luggage scanning device.

In identity verification system, the threshold value setting for the image similarity is very vital. If the threshold value is too high, many eligible people could not pass the identity verification and the false alarm occurs too frequently. Conversely, if the threshold value is too low, those unqualified people will pass the identity verification. In this case, the system is invalid. So, the setting of threshold value should keep balance between the rapidness and the false alarm rate. The trade-off acquisition is handcrafted and time-consuming. There are no specific rules or principles to guide the setting, just depend on the practical experiments in particular cases. In our further work, we will study the technique for setting optimum threshold value.

In the following paragraphs, we will introduce the two core techniques in the image similarity-based identity verification system. One is the super-

131

resolution reconstruction, the other is the deep learning-based face detection.

Generally speaking, the super resolution construction are some kinds of restoration techniques, which consists frequency domain algorithm and time domain algorithm, for the original high resolution image based on multi-frame low resolution images (Zhang, 2010). All the low resolution images is captured in the same scene with the original high resolution image and there just exists slight changes. If there only exists one low resolution, the ordinary method to get the high resolution image is interpolation.

In the case of only one low resolution image, different form the traditional interpolation method, in (Luo, 2011) the authors proposed the deep learning-based strategy for single image super-resolution. With light weighted structure deep convolution neural network (CNN), this method directed learns an end-to-end mapping between the low/high resolution images. They also proved that the sparse-coding-based SR can be viewed as a convolutional neural network. This work claimed the state-of-the-art performance and suitable for the online usage.

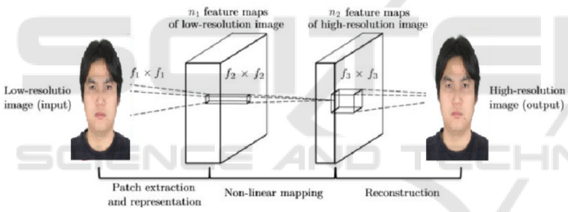

Figure 4: Given a low resolution image Y, the first convolution layer extracts a set of feature maps. The second layer maps these feature maps nonlinearly to high resolution patch representation. The last layer combines the predictions with a spatial neighbourhood to produce the final high resolution image F(Y).

In this above-mentioned work, the authors took the low resolution image as the input and output the high resolution one. To execute the image quality enhancement using this deep-learning-based method, the training stage should be carried out prior to the output stage. Refer to this method, we utilize over 5000 pairs of LR images and HR images, which with 126*102 pixels and 441*358 pixels respectively, as training dataset.

Fig.4 shows the schematic diagram of the deep learning-based SR.

Face detection in the complex scenes is an essential but rarely rough task. To the fixed surveillance camera, the field of view (FOV) is constant. In this scene, the face region in the frame image is enough bit to execute the face detection. But in the ordinary surveillance scenes, to those peoples far away from the fixed-focus camera, the face region maybe too small to be detected. In this case, the pedestrian detection should be utilized to detect the concerned people and track this people until his approach makes the face region enough big to be detected. This strategy was proposed in our previous work (Yan, 2014) and proved to be effective and efficient.

Considering the complexity of face detection in the ordinary surveillance scenes, the researchers presents a new state-of-the art approach in (Chen, 2014). They observed that the aligned face shapes provides better features for face classification. To combine the face alignment and detection more effectively, they learned this two tasks in the same cascade. By exploiting the joint learning, the capability of cascade detection and real time performance can both achieve the satisfied status.

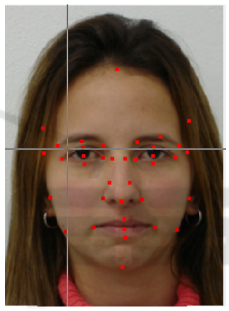

Figure 5: The key point annotation on face shape.

As shown in Fig.5, we use 38 key points to describe the face shape, 10 points for face contour, 6 points for eyebrows, 10 points for eyes, 7 points for nose and 5 points for lip respectively.

We bought a face image dataset consisted of about 20, 000 face images and 20, 000 natural scene images without faces from web. All face images are transferred into grayscale images. After all the face images are labelled, the dataset is utilized to train the classification/regression tree.

3 EXPERIMENT AND CONCLUSION

We utilized the combination of the on-site surveillance camera and RFID reader to realize the self-service passenger pass. The key techniques focus in the on-site face detection effectively and the online SR reconstruction for the low resolution ID electronic photos. As a comparison, we also directly

132

The Study on the Identity Verification between the on-Site Face Image and the ID Photo

adopted the electronic ID photos without the SR construction.

The subject consists of 131 peoples. In the experiments, 5 people face can't be detected successfully. Among the other 126 peoples, 102 peoples can pass the security check with the online SR construction, nearly 80.9% hit rate. At the same scene, without the SR construction, only 46 peoples among the 126 peoples can pass the security check gate with self-service, nearly 36.5% hit rate. The interface of the identity verification system is shown in fig. 6.

Figure 6: The interface of the Identity Verification System.

As mentioned before, some face image can't be successfully detected in the surveillance vision. This default is partly due to the limited FOV of the focus-fixed camera. In our future work, we will adopt the dual-camera configuration consists of a static camera and an active camera to replace the single focus-fixed camera, see fig.7.

In fig.7, the static camera is a fixed-focus camera with wide view range and in charge of the pedestrian detection and transmit the corresponding position information to the active camera which has the variable focus. The active camera and track the pedestrian and grab the clear HR face image for identity verification system.

Figure 7: The dual-camera configuration in Identity verification system.

The proposed identity verification system is based on the image similarity measurement, and the practical experiments show its effective in practice. On the other hand, frankly speaking, the performance without the SR construction is inferior to the expectation, i.e., the electronic ID photo is unsuitable to be used directly for identity verification.

ACKNOWLEDGEMENTS

Our research was supported by the Innovation plan for practical application of the Ministry of Public Security (No. 2012YYCXGASS125), National Science and Technology Support Projects of China (No. 2012BAH07B01), the Project of Shanghai Municipal Commission Of Economy and Information (No. 12DZ0512100)，Major program of Science and Technology Commission of Shanghai Municipality (No.12510701900), National High-tech R&D Program of China “863 Program ” (No. 2013AA01A603).

REFERENCES

Dong, C., et al., *Image Super-Resolution Using Deep Convolution Networks.* arXiv preprint arXiv:1501.00092, 2014: p. 12.

Yan, Z., F. Yang, and J. Wang, *Face Orientation Detection in Video Stream based on Haar-Like Feature and LQV Classifier for Civil Video Surveillance*, in *2013 IET Second International Conference on Smart and Sustainable City*2013: Zhangjiajie, Hunan. p. 5.

Zhang, L., H. Zhang, and H. Shen, *A super-resolution reconstruction algorithm for surveillance images.* Singal Processing, 2010(90): p. 12.

Yan, Z. and K. Gu, *The Study on the Techniques on Gun-Dome Camera Cooperative Human-Tracing.* The 10th International conference on Semantic, Knowledge and Grid, 2014: p. 6.

Chen, D., S. Ren, and Y. Wei, *Joint Cascade Face Detection and Alignment.* European Conference on Computer Vision (ECCV),, 2014: p. 14.

Luo, X., Xu, Z., Yu, J., and Chen, X. 2011. Building Association Link Network for Semantic Link on Web Resources. *IEEE transactions on automation science and engineering*, 8(3):482-494.

133